



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/712,445	11/13/2003	Samuel L. Karns	BOC920030105US1 (002)	8692
46322 7590 11/12/2008 CAREY, RODRIGUEZ, GREENBERG & PAUL, LLP STEVEN M. GREENBERG 950 PENINSULA CORPORATE CIRCLE SUITE 3020 BOCA RATON, FL 33487				
EXAMINER YEN, ERIC L.				
ART UNIT 2626		PAPER NUMBER		
MAIL DATE 11/12/2008		DELIVERY MODE PAPER		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary

Application No.

10/712,445

Applicant(s)

KARNS, SAMUEL L.

Examiner

ERIC YEN

Art Unit

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 30 July 2008.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-20 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-20 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO/SI/08)
Paper No(s)/Mail Date _____
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date _____
- 5) ☐ Notice of Informal Patent Application
- 6) ☐ Other: _____

DETAILED ACTION

Response to Amendment

1. In response to the Office Action mailed 4/30/08, applicant has submitted an amendment filed 7/30/08.

Arguments for allowability have been presented.

Response to Arguments

2. Applicant's arguments with respect to claims 1-20 have been considered but are moot in view of the new ground(s) of rejection.
3. Applicant's arguments filed 7/30/08 have been fully considered but they are not persuasive.

Applicant argues that "language phoneme data" should be given the "plain meaning" of "the plurality phonemes that occur in a given language" (Amendment, page 13).

The examiner respectfully disagrees, because that is not the only meaning that can be attributed to "language phoneme data". An alternative meaning that is still "plain" is data pertaining to the language's phonemes. The reference corpus includes data pertaining to how the language's phonemes are used (i.e., the sequence of phonemes that make up specific words) and so is "language phoneme data" in that sense. If applicant wants to import the definition "data limited to the plurality of

phonemes that occur in a given language" it should be amended into the claims to prevent alternative interpretations from reading on the claims under the principle of giving claims their broadest reasonable interpretation.

Applicant then asserts that the examiner is equating selection of sentences from a "corpus of sentences" with the script data in claims 1 and 10 and believes that the examiner is mapping two separate and distinct claim limitations, specifically "language phoneme data" and "script data" to be the same thing, which is a sample of text, and argues that "the examiner cannot properly ascribe the same meaning to two separate and distinct claim limitations" (Amendment, page 14). Applicant notes that the "corpus of sentences" is simply a subset of the reference corpus of Esquerra" and "thus, one item, the former, is simply a portion or subset of the latter" (Amendment, page 14).

Applicant appears to be arguing that just because the nature of two items is the same (i.e., that they are samples of text from the same source) that this necessarily means that the limitations are mapped to the same thing. This is not true, for the same reason that a portion of a whole is not the whole itself, even if the portion and the whole contain the same substance. If one limitation is mapped to the whole and another is mapped to the portion, then it is certainly not true that the limitation mapped to the portion is being mapped to the same thing because the portion excludes other parts of the whole. So, in this case, the subset mapped to the "script data" is not the same as the entire reference corpus mapped to "language phoneme data".

Applicant then argues that the definition of a script is "text that is read aloud by a user so as to garner an example of a particular user's voice signature and speaking style" and so the examiner's assertion that the "selection of sentences from the 'corpus of sentences' are scripts because they are a collection of texts from which phoneme data can be retrieved" is inaccurate (Amendment, page 14).

Since this is specifically defined in the Specification, the definition is to be imported into the claims.

However, this still does not allow the invention to be patentable. Applicant's Specification does not only include this definition of "script", but also includes the admission that this type of script as speech recognition data is common and, therefore, it is not novel. Specifically, this definition and the use of scripts appears in applicant's background and describes that this is not something new because it is already common. Therefore, this portion of applicant's Specification qualifies as admitted prior art.

Applicant then argues that "allophones are not the same phonemes since one is a variation of the other", and "because an allophone is one phone of many that belong to the same phoneme, it goes to stand that counting allophones does not provide an accurate count of phoneme[s]", and that "rather, after all allophones are counted, a separate categorization of allophones into corresponding phonemes must be made in order to reach an accurate count of phonemes" (Amendment, page 16).

Applicant appears to be misunderstanding what the examiner's application of the art is. Even if counting one allophone belonging to a phoneme does not yield the total

count for all variations of one phoneme, a count for each allophone is still one count for the phoneme since the allophone, as a variation, is a narrower category of a phoneme, but is still the phoneme itself. Applicant seems to recognize this based on the argument that other allophones must be counted to yield a total for the phoneme. The claim language, however, recites "counting each phoneme in the script data to produce a count data for each of the plurality of phonemes in the language phoneme data". If the count of the allophones is a count of a subset of all the phonemes that appear in a corpus, then it is still a count for a particular phoneme because the allophone is the phoneme itself in a particular context. Since the allophone is the phoneme itself, a count of an allophone is a count of a phoneme. Applicant's claim language does not say that the counting produces "the total number of appearances of a phoneme in the language phoneme data", and so the scope of a count for a phoneme is not limited to what applicant is arguing that the count data is.

Applicant then asserts that the reference is not enabling (Amendment, pages 16-17), but at the very least the exact quotes in applicant's remarks are not applicable because the examiner did not say the Esquerre anticipated the claims, but rather rendered it obvious. Esquerre makes this obvious to one of ordinary skill in the art by pointing out the source of the corpus is the internet, and also the counts and processing would be impractical for a person to do manually since corpora generally have a large quantity of content. Therefore, the inference one of ordinary skill in the art would make is that some sort of computer program can be designed to do the counting. Generally,

specifications are not in such great detail that every specific detail pertaining to a particular step necessary to implement a particular method step, for example.

Therefore, it would be unreasonable to think that the standard for enabling disclosure, even if it is required for a showing of obviousness (in contrast with anticipation) is that every specific detail must be set forth in the reference's description. So, given the acknowledged existence of the internet in Esquerra and the fact that computers connected to the internet and used for data processing did exist in 1998 when Esquerra was published, among other things, one of ordinary skill in the art would be able to make and use the invention based on Esquerra's description of what the computer program would have to do.

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

1. **Claims 1, 4-6, 10, 13-15 and 19** are rejected under 35 U.S.C. 103(a) as being unpatentable over ESQUERRA et al. ("Design of a Phonetic Corpus for Speech Recognition in Catalan"), in view of Applicant's Admitted Prior Art, hereafter AAPA.

2. Regarding **claim 1**, ESQUERRA teaches a method for developing a script ("corpus of sentences", section 3.1) to be used with speech recognition systems ("for speech recognition", abstract), said method comprising the steps of: reading language phoneme data ("reference corpus", section 2) for a given language, the language phoneme data having a plurality of phonemes occurring in the given language ("corpus was converted into phonemes using a transcription program", section 2.1);

reading script data ("sentences between 10 and 40 letters were selected", section 3.1) having a set of one or more phonemes ("N is the number of phones in a sentence", section 3.1; "text-to-phoneme", Section 2.1; see Response to Arguments);

counting each phoneme in the script data to produce a count data for each of the plurality of phonemes in the language phoneme data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, "text-to-phoneme", Section 2.1; where "units" refer to phonemes); generating a set of statistical data ("coverage measures", section 4, paragraph 5) derived from the count data, the set of statistical data including one or more metrics of the extent to which the phonemes in the language phoneme data are included in the script data (see Table 3, BD3-E is the corpus of sentences used for training, REF is the reference corpus).

Esquerra fails to teach a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style.

AAPA teaches a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style (Specification, paragraph 3).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Esquerria to include the teaching of AAPA of a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style, in order to improve recognition performance, as described by Newman et al. (US 6,151,575), hereafter Newman (col. 1, lines 50-65; "classify different phonemes... supervised... having the speaker read from a script", col. 8, lines 24-45)

3. Regarding **claim 4**, ESQUERRA further teaches that the set of statistical data includes:

an occurrence data for each of the phonemes in the phoneme data, each occurrence data indicating a number of occurrences of the phoneme in the script data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, paragraph 1, where "units" refer to phonemes).

4. Regarding **claim 5**, ESQUERRA further teaches that the set of statistical data includes:

a ratio data, each ratio data being the number of phonemes in the script data as a percentage of the number of the plurality of phonemes in the phoneme data (see

Table 3, BD3-E is the corpus of sentences used for training, REF is the reference corpus).

5. Regarding **claim 6**, ESQUERRA further teaches that the set of statistical data includes:

a missing phoneme data, each missing phoneme data being a list of the phonemes in the language phoneme data not included in the script data (see section 3.1, paragraph 2, new sentences are created containing missing allophones, so a list of the missing allophones is inherent).

6. Regarding **claim 10**, ESQUERRA teaches a machine readable storage having stored thereon a computer program for developing a script ("corpus of sentences", section 3.1, paragraph 1; "Internet", Section 2; See Response to Arguments) to be used with speech recognition systems ("for speech recognition", abstract), said computer program comprising a routine set of instructions for causing the machine to perform the steps of:

reading language phoneme data ("reference corpus", section 2) for a given language, the language phoneme data having a plurality of phonemes occurring in the given language ("corpus was converted into phonemes using a transcription program", section 2.);

reading script data ("sentences between 10 and 40 letters were selected", section 3.1; "text-to-phoneme", Section 2.1) having a set of one or more phonemes ("N is the number of phones in a sentence", section 3.1);

counting each phoneme in the script data to produce a count data for each of the plurality of phonemes in the language phoneme data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1; "text-to-phoneme", Section 2.1, Table 1; where "units" refer to phonemes); generating a set of statistical data ("coverage measures", section 4, paragraph 5) derived from the count data, the set of statistical data including one or more metrics of the extent to which the phonemes in the language phoneme data are included in the script data (see Table 3, BD3-E is the corpus of sentences used for training, REF is the reference corpus).

Esquerra fails to teach a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style.

AAPA teaches a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style (Specification, paragraph 3).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Esquerra to include the teaching of AAPA of a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style, in order to improve recognition performance, as described by

Newman et al. (US 6,151,575), hereafter Newman (col. 1, lines 50-65; "classify different phonemes... supervised... having the speaker read from a script", col. 8, lines 24-45)

7. Regarding **claim 13**, ESQUERRA further teaches that the set of statistical data includes:

an occurrence data for each of the phonemes in the phoneme data, each occurrence data indicating a number of occurrences of the phoneme in the script data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, paragraph 1, where "units" refer to phonemes).

8. Regarding **claim 14**, ESQUERRA further teaches that the set of statistical data includes:

a ratio data, each ratio data being the number of phonemes in the script data as a percentage of the number of the plurality of phonemes in the phoneme data (see Table 3, BD3-E is the corpus of sentences used for training, REF is the reference corpus).

9. Regarding **claim 15**, ESQUERRA further teaches that the set of statistical data includes:

a missing phoneme data, each missing phoneme data being a list of the phonemes in the language phoneme data not included in the script data (see section

3.1, paragraph 2, new sentences are created containing missing allophones, so a list of the missing allophones is inherent).

10. Regarding **claim 19**, ESQUERRA teaches a script development tool ("design of a corpus for speech recognition", abstract) configured for coupling to a script ("corpus of sentences", section 3.1) having a set of one or more phonemes ("N is the number of phones in a sentence", section 3.1; "text-to-phoneme", Section 2.1) and programmed to both count each phoneme in said script to produce count data for each phoneme in a selected language ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, Table 1; "text-to-phoneme", Section 2.1; where "units" refer to phonemes), and also to generate a set of statistical data ("coverage measures", section 4, paragraph 5) derived from said count data, the set of statistical data comprising one or more metrics of the extent to which each phoneme in said selected language is included in said script (see Table 3, BD3-E is the corpus of sentences used for training, REF is the reference corpus).

Esquerra fails to teach a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style.

AAPA teaches a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style (Specification, paragraph 3).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Esquerra to include the teaching of AAPA of a script is something

that is read aloud by the end user as an example of a particular users' voice signature and speaking style, in order to improve recognition performance, as described by Newman et al. (US 6,151,575), hereafter Newman (col. 1, lines 50-65; "classify different phonemes... supervised... having the speaker read from a script", col. 8, lines 24-45)

11. **Claims 2, 3, 7, 8, 11, 12, 16, 17, and 20** are rejected under 35 U.S.C. 103(a) as being unpatentable over ESQUERRA et al. ("Design of a Phonetic Corpus for Speech Recognition in Catalan") in view of GOULD (Patent No.: US 5,794,189).

12. Regarding **claim 2**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), each word having one or more of the set of one or more phonemes ("N is the number of phones in a sentence", section 3.1, paragraph 1).

However, ESQUERRA, in view of AAPA, does not disclose reading vocabulary data, comparing words to vocabulary data, or returning an error message.

In the same field of speech recognition, GOULD teaches:

reading vocabulary data having one or more words ("dictionary", column 15, line 15);

comparing each word in the script data with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", column 15, lines 14-15); and

returning an error message if a word in the script data is not included in the vocabulary data ("if the word is not in the dictionary... display an 'unknown word' error to the user", column 15, lines 16-20).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the corpus of sentences provided by ESQUERRA, in view of AAPA, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

Regarding **claim 3**, ESQUERRA teaches counting each phoneme in each word in the script data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, paragraph 1, where "units" refer to phonemes).

However, ESQUERRA, in view of AAPA, does not disclose comparing words to vocabulary data, returning an error message, or counting the phonemes if the word is in the vocabulary data.

In the same field of speech recognition, ESQUERRA in view of GOULD teach:
comparing each word in the script data with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", GOULD, column 15, lines 14-15);
returning an error message if a word in the script data is not included in the vocabulary data ("if the word is not in the dictionary... display an 'unknown word' error to the user", GOULD, column 15, lines 16-20); and

counting each phoneme in each word in the script data ("units were counted to know whether they reach the minimum number of required repetitions", ESQUERRA, section 3.1, paragraph 1) if a word in the script data is included in the vocabulary data ("remember these words as target words", column 15, line 21, where a word marked as a target word has further operations performed on it).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the corpus of sentences provided by ESQUERRA, in view of AAPA, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

13. Regarding **claim 7**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), and further comprising the steps of:

reading an additional word having one or more phonemes ("new sentences had to be written containing those allophones", section 3.1, paragraph 2); and

adding an additional word to the script data ("new sentences were added to the corpus", section 3.1, paragraph 4).

However ESQUERRA, in view of AAPA, does not disclose reading a vocabulary data, comparing the additional word to the vocabulary data, or adding the additional word if the additional word is included in the vocabulary data.

In the same field of speech recognition, ESQUERRA in view of GOULD teach:
reading a vocabulary data having one or more words ("dictionary", GOULD, column 15, line 15);

comparing the additional word with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", GOULD, column 15, lines 14-15);

adding the additional word to the script data ("new sentences were added to the corpus", section 3.1, paragraph 4) if the additional word is included in the vocabulary data ("remember these words as target words", column 15, line 21, where a word marked as a target word has further operations performed on it).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the new sentence provided by ESQUERRA, in view of AAPA, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

14. Regarding **claim 8**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), and further comprising the step of:

reading an additional word having one or more phonemes ("new sentences had to be written containing those allophones", section 3.1, paragraph 2);

However ESQUERRA, in view of AAPA, does not disclose reading a vocabulary data, comparing the additional word with the script data, or removing the additional word from the script data.

In the same field of speech recognition, GOULD teaches:

reading a vocabulary data having one or more words ("dictionary", column 15, line 15);

comparing the additional word with the script data ("if the text on the screen starting with the current word matches the indicated words, set the selection to text on the screen just compared against", column 13, lines 35-38, where "text on the screen" is the additional word, and the "indicated words" is the script data);

removing the additional word from the script data if the additional word is included in the script data ("if words are selected on the screen, delete the words which are selected", column 13, lines 48-49, where the word "selected on the screen" is the additional word that was compared with the script data).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to delete the words in the new sentence provided by ESQUERRA, in view of AAPA, in the manner of GOULD in order to filter out words which may be problematic for training (see PITRELLI et al., section 2.1, listed on form PTO-892).

15. Regarding **claim 11**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between

10 and 40 letters were selected”, section 3.1, paragraph 1), each word having one or more of the set of one or more phonemes (“N is the number of phones in a sentence”, section 3.1, paragraph 1).

However, ESQUERRA, in view of AAPA, does not disclose reading vocabulary data, comparing words to vocabulary data, or returning an error message.

In the same field of speech recognition, GOULD teaches:

reading vocabulary data having one or more words (“dictionary”, column 15, line 15);

comparing each word in the script data with the vocabulary data (“for each word in the buffer, look the word up in the dictionary”, column 15, lines 14-15); and

returning an error message if a word in the script data is not included in the vocabulary data (“if the word is not in the dictionary... display an ‘unknown word’ error to the user”, column 15, lines 16-20).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the corpus of sentences provided by ESQUERRA, in view of AAPA, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

Regarding **claim 12**, ESQUERRA teaches counting each phoneme in each word in the script data (“units were counted to know whether they reach the minimum number of required repetitions”, section 3.1, paragraph 1, where “units” refer to phonemes).

However, ESQUERRA, in view of AAPA, does not disclose comparing words to vocabulary data, returning an error message, or counting the phonemes if the word is in the vocabulary data.

In the same field of speech recognition, ESQUERRA in view of GOULD teach:
comparing each word in the script data with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", GOULD, column 15, lines 14-15);
returning an error message if a word in the script data is not included in the vocabulary data ("if the word is not in the dictionary... display an 'unknown word' error to the user", GOULD, column 15, lines 16-20); and

counting each phoneme in each word in the script data ("units were counted to know whether they reach the minimum number of required repetitions", ESQUERRA, section 3.1, paragraph 1) if a word in the script data is included in the vocabulary data ("remember these words as target words", column 15, line 21, where a word marked as a target word has further operations performed on it).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the corpus of sentences provided by ESQUERRA, in view of AAPA, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

16. Regarding **claim 16**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between

10 and 40 letters were selected", section 3.1, paragraph 1), and further comprising the steps of:

reading an additional word having one or more phonemes ("new sentences had to be written containing those allophones", section 3.1, paragraph 2); and

adding an additional word to the script data ("new sentences were added to the corpus", section 3.1, paragraph 4).

However ESQUERRA, in view of AAPA, does not disclose reading a vocabulary data, comparing the additional word to the vocabulary data, or adding the additional word if the additional word is included in the vocabulary data.

In the same field of speech recognition, ESQUERRA in view of GOULD teach: reading a vocabulary data having one or more words ("dictionary", GOULD, column 15, line 15);

comparing the additional word with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", GOULD, column 15, lines 14-15);

adding the additional word to the script data ("new sentences were added to the corpus", ESQUERRA, section 3.1, paragraph 4) if the additional word is included in the vocabulary data ("remember these words as target words", column 15, line 21, where a word marked as a target word has further operations performed on it).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the new sentence provided by ESQUERRA, in view of AAPA, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

17. Regarding **claim 17**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), and further comprising the step of:

reading an additional word having one or more phonemes ("new sentences had to be written containing those allophones", section 3.1, paragraph 2);

However ESQUERRA, in view of AAPA, does not disclose reading a vocabulary data, comparing the additional word with the script data, or removing the additional word from the script data.

In the same field of speech recognition, GOULD teaches:

reading a vocabulary data having one or more words ("dictionary", column 15, line 15);

comparing the additional word with the script data ("if the text on the screen starting with the current word matches the indicated words, set the selection to text on the screen just compared against", column 13, lines 35-38, where "text on the screen" is the additional word, and the "indicated words" is the script data);

removing the additional word from the script data if the additional word is included in the script data ("if words are selected on the screen, delete the words which are selected", column 13, lines 48-49, where the word "selected on the screen" is the additional word that was compared with the script data).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to delete the words in the new sentence provided by ESQUERRA, in view of AAPA, in the manner of GOULD in order to filter out words which may be problematic for training (see PITRELLI et al., section 2.1, listed on form PTO-892).

18. Regarding **claim 20**, ESQUERRA teaches that the script ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), and wherein the tool is further programmed to read an additional word having one or more phonemes ("new sentences had to be written containing those allophones", section 3.1, paragraph 2), and add the additional word to the script data ("new sentences were added to the corpus", section 3.1, paragraph 4).

However ESQUERRA, in view of AAPA, does not disclose a tool that is programmed to read a vocabulary data having one or more words, and is also programmed to compare the additional word with the vocabulary data and add the additional word to the script data if the additional word is included in the vocabulary data, and is also programmed to compare the additional word with the script and remove the additional word from the script data if the additional word is included in the script data.

In the same field of speech recognition, ESQUERRA in view of GOULD teach a tool that is programmed to read a vocabulary data having one or more words

("dictionary", GOULD, column 15, line 15), and is also programmed to compare the additional word with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", GOULD, column 15, lines 14-15) and add the additional word to the script data ("new sentences were added to the corpus", ESQUERRA, section 3.1, paragraph 4) if the additional word is included in the vocabulary data ("remember these words as target words", column 15, line 21, where a word marked as a target word has further operations performed on it), and is also programmed to compare the additional word with the script ("if the text on the screen starting with the current word matches the indicated words, set the selection to text on the screen just compared against", GOULD, column 13, lines 35-38, where "text on the screen" is the additional word, and the "indicated words" is the script data) and remove the additional word from the script data if the additional word is included in the script data ("if words are selected on the screen, delete the words which are selected", GOULD, column 13, lines 48-49, where the word "selected on the screen" is the additional word that was compared with the script data).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the new sentence provided by ESQUERRA, in view of AAPA, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20) and to filter out words which may be problematic for training (see PITRELLI et al., section 2.1, listed on form PTO-892).

19. **Claims 9 and 18** are rejected under 35 U.S.C. 103(a) as being unpatentable over ESQUERRA et al. ("Design of a Phonetic Corpus for Speech Recognition in Catalan"), in view of AAPA, and further in view of Department of Psychology, University of Essex ("Phoneme Search"), hereinafter referred to as ESSEX.

20. Regarding **claim 9**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1).

However, ESQUERRA, in view of AAPA, does not disclose reading a vocabulary data, reading a set of one or more desired phonemes, searching the vocabulary data for one or more words having the set of one or more desired phonemes, or generating a report of one or more additional words having the set of one or more desired phonemes.

In the same field of phonetic evaluation, ESSEX teaches:

reading a vocabulary data having one or more words ("word database", see header);

reading a set of one or more desired phonemes (three different phonemes may be selected with the pull-down menus);

searching the vocabulary data for one or more words having the set of one or more desired phonemes ("search for words which contain the following phonemes");

generating a report of one or more additional words having the set of one or more desired phonemes ("generates a list of words"), if the one or more additional

words having the set of one or more desired phonemes are included in the vocabulary data (see "Phoneme Search Results").

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made use the phoneme search engine of ESSEX with the corpus design of ESQUERRA, in view of AAPA, in order to find words containing "missing units" (ESSEX, section 5, paragraph 1).

21. Regarding **claim 18**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1).

However, ESQUERRA, in view of AAPA, does not disclose reading a vocabulary data, reading a set of one or more desired phonemes, searching the vocabulary data for one or more words having the set of one or more desired phonemes, or generating a report of one or more additional words having the set of one or more desired phonemes.

In the same field of phonetic evaluation, ESSEX teaches:

reading a vocabulary data having one or more words ("word database", see header);

reading a set of one or more desired phonemes (three different phonemes may be selected with the pull-down menus);

searching the vocabulary data for one or more words having the set of one or more desired phonemes ("search for words which contain the following phonemes");

generating a report of one or more additional words having the set of one or more desired phonemes ("generates a list of words"), if the one or more additional words having the set of one or more desired phonemes are included in the vocabulary data (see "Phoneme Search Results").

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made use the phoneme search engine of ESSEX with the corpus design of ESQUERRA, in view of AAPA, in order to find words containing "missing units" (ESSEX, section 5, paragraph 1).

22. **Claims 1, 4-6, 10, 13-15 and 19** are rejected under 35 U.S.C. 103(a) as being unpatentable over ESQUERRA et al. ("Design of a Phonetic Corpus for Speech Recognition in Catalan"), in view of Applicant's Admitted Prior Art, hereafter NEWMAN.

23. Regarding **claim 1**, ESQUERRA teaches a method for developing a script ("corpus of sentences", section 3.1) to be used with speech recognition systems ("for speech recognition", abstract), said method comprising the steps of: reading language phoneme data ("reference corpus", section 2) for a given language, the language phoneme data having a plurality of phonemes occurring in the given language ("corpus was converted into phonemes using a transcription program", section 2.1);

reading script data ("sentences between 10 and 40 letters were selected", section 3.1) having a set of one or more phonemes ("N is the number of phones in a sentence", section 3.1; "text-to-phoneme", Section 2.1; see Response to Arguments);

counting each phoneme in the script data to produce a count data for each of the plurality of phonemes in the language phoneme data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, "text-to-phoneme", Section 2.1; where "units" refer to phonemes);
generating a set of statistical data ("coverage measures", section 4, paragraph 5) derived from the count data, the set of statistical data including one or more metrics of the extent to which the phonemes in the language phoneme data are included in the script data (see Table 3, BD3-E is the corpus of sentences used for training, REF is the reference corpus).

Esquerria fails to teach a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style.

NEWMAN teaches a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style (col. 1, lines 50-65; "classify different phonemes... supervised... having the speaker read from a script", col. 8, lines 24-45).

Therefore, it would have been obvious to one of ordinary skill in the art a the time of invention to modify Esquerria to include the teaching of NEWMAN of a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style, in order to improve recognition performance, as described

by Newman et al. (US 6,151,575), hereafter Newman (col. 1, lines 50-65; "classify different phonemes... supervised... having the speaker read from a script", col. 8, lines 24-45)

24. Regarding **claim 4**, ESQUERRA further teaches that the set of statistical data includes:

an occurrence data for each of the phonemes in the phoneme data, each occurrence data indicating a number of occurrences of the phoneme in the script data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, paragraph 1, where "units" refer to phonemes).

25. Regarding **claim 5**, ESQUERRA further teaches that the set of statistical data includes:

a ratio data, each ratio data being the number of phonemes in the script data as a percentage of the number of the plurality of phonemes in the phoneme data (see Table 3, BD3-E is the corpus of sentences used for training, REF is the reference corpus).

26. Regarding **claim 6**, ESQUERRA further teaches that the set of statistical data includes:

a missing phoneme data, each missing phoneme data being a list of the phonemes in the language phoneme data not included in the script data (see section

3.1, paragraph 2, new sentences are created containing missing allophones, so a list of the missing allophones is inherent).

27. Regarding **claim 10**, ESQUERRA teaches a machine readable storage having stored thereon a computer program for developing a script ("corpus of sentences", section 3.1, paragraph 1; "Internet", Section 2; See Response to Arguments) to be used with speech recognition systems ("for speech recognition", abstract), said computer program comprising a routine set of instructions for causing the machine to perform the steps of:

reading language phoneme data ("reference corpus", section 2) for a given language, the language phoneme data having a plurality of phonemes occurring in the given language ("corpus was converted into phonemes using a transcription program", section 2.);

reading script data ("sentences between 10 and 40 letters were selected", section 3.1; "text-to-phoneme", Section 2.1) having a set of one or more phonemes ("N is the number of phones in a sentence", section 3.1);

counting each phoneme in the script data to produce a count data for each of the plurality of phonemes in the language phoneme data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1; "text-to-phoneme", Section 2.1, Table 1; where "units" refer to phonemes);
generating a set of statistical data ("coverage measures", section 4, paragraph 5) derived from the count data, the set of statistical data including one or more metrics of

the extent to which the phonemes in the language phoneme data are included in the script data (see Table 3, BD3-E is the corpus of sentences used for training, REF is the reference corpus).

Esquerra fails to teach a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style.

NEWMAN teaches a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style (col. 1, lines 50-65; "classify different phonemes... supervised... having the speaker read from a script", col. 8, lines 24-45).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Esquerra to include the teaching of NEWMAN of a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style, in order to improve recognition performance, as described by Newman et al. (US 6,151,575), hereafter Newman (col. 1, lines 50-65; "classify different phonemes... supervised... having the speaker read from a script", col. 8, lines 24-45)

28. Regarding **claim 13**, ESQUERRA further teaches that the set of statistical data includes:

an occurrence data for each of the phonemes in the phoneme data, each occurrence data indicating a number of occurrences of the phoneme in the script data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, paragraph 1, where "units" refer to phonemes).

29. Regarding **claim 14**, ESQUERRA further teaches that the set of statistical data includes:

a ratio data, each ratio data being the number of phonemes in the script data as a percentage of the number of the plurality of phonemes in the phoneme data (see Table 3, BD3-E is the corpus of sentences used for training, REF is the reference corpus).

30. Regarding **claim 15**, ESQUERRA further teaches that the set of statistical data includes:

a missing phoneme data, each missing phoneme data being a list of the phonemes in the language phoneme data not included in the script data (see section 3.1, paragraph 2, new sentences are created containing missing allophones, so a list of the missing allophones is inherent).

31. Regarding **claim 19**, ESQUERRA teaches a script development tool ("design of a corpus for speech recognition", abstract) configured for coupling to a script ("corpus of sentences", section 3.1) having a set of one or more phonemes ("N is the number of

phones in a sentence", section 3.1; "text-to-phoneme", Section 2.1) and programmed to both count each phoneme in said script to produce count data for each phoneme in a selected language ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, Table 1; "text-to-phoneme", Section 2.1; where "units" refer to phonemes), and also to generate a set of statistical data ("coverage measures", section 4, paragraph 5) derived from said count data, the set of statistical data comprising one or more metrics of the extent to which each phoneme in said selected language is included in said script (see Table 3, BD3-E is the corpus of sentences used for training, REF is the reference corpus).

Esquerra fails to teach a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style.

NEWMAN teaches a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style (col. 1, lines 50-65; "classify different phonemes... supervised... having the speaker read from a script", col. 8, lines 24-45).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Esquerra to include the teaching of NEWMAN of a script is something that is read aloud by the end user as an example of a particular users' voice signature and speaking style, in order to improve recognition performance, as described by Newman et al. (US 6,151,575), hereafter Newman (col. 1, lines 50-65; "classify different phonemes... supervised... having the speaker read from a script", col. 8, lines 24-45)

32. **Claims 2, 3, 7, 8, 11, 12, 16, 17, and 20** are rejected under 35 U.S.C. 103(a) as being unpatentable over ESQUERRA et al. ("Design of a Phonetic Corpus for Speech Recognition in Catalan") in view of GOULD (Patent No.: US 5,794,189).

33. Regarding **claim 2**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), each word having one or more of the set of one or more phonemes ("N is the number of phones in a sentence", section 3.1, paragraph 1).

However, ESQUERRA, in view of NEWMAN, does not disclose reading vocabulary data, comparing words to vocabulary data, or returning an error message.

In the same field of speech recognition, GOULD teaches:

reading vocabulary data having one or more words ("dictionary", column 15, line 15);

comparing each word in the script data with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", column 15, lines 14-15); and

returning an error message if a word in the script data is not included in the vocabulary data ("if the word is not in the dictionary... display an 'unknown word' error to the user", column 15, lines 16-20).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the corpus of sentences

provided by ESQUERRA, in view of NEWMAN, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

Regarding **claim 3**, ESQUERRA teaches counting each phoneme in each word in the script data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, paragraph 1, where "units" refer to phonemes).

However, ESQUERRA, in view of NEWMAN, does not disclose comparing words to vocabulary data, returning an error message, or counting the phonemes if the word is in the vocabulary data.

In the same field of speech recognition, ESQUERRA in view of GOULD teach:
comparing each word in the script data with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", GOULD, column 15, lines 14-15);
returning an error message if a word in the script data is not included in the vocabulary data ("if the word is not in the dictionary... display an 'unknown word' error to the user", GOULD, column 15, lines 16-20); and

counting each phoneme in each word in the script data ("units were counted to know whether they reach the minimum number of required repetitions", ESQUERRA, section 3.1, paragraph 1) if a word in the script data is included in the vocabulary data ("remember these words as target words", column 15, line 21, where a word marked as a target word has further operations performed on it).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the corpus of sentences provided by ESQUERRA, in view of NEWMAN, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

34. Regarding **claim 7**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), and further comprising the steps of:

reading an additional word having one or more phonemes ("new sentences had to be written containing those allophones", section 3.1, paragraph 2); and

adding an additional word to the script data ("new sentences were added to the corpus", section 3.1, paragraph 4).

However ESQUERRA, in view of NEWMAN, does not disclose reading a vocabulary data, comparing the additional word to the vocabulary data, or adding the additional word if the additional word is included in the vocabulary data.

In the same field of speech recognition, ESQUERRA in view of GOULD teach:
reading a vocabulary data having one or more words ("dictionary", GOULD, column 15, line 15);

comparing the additional word with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", GOULD, column 15, lines 14-15);

adding the additional word to the script data ("new sentences were added to the corpus", section 3.1, paragraph 4) if the additional word is included in the vocabulary data ("remember these words as target words", column 15, line 21, where a word marked as a target word has further operations performed on it).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the new sentence provided by ESQUERRA, in view of NEWMAN, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

35. Regarding **claim 8**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), and further comprising the step of:

reading an additional word having one or more phonemes ("new sentences had to be written containing those allophones", section 3.1, paragraph 2);

However ESQUERRA, in view of NEWMAN, does not disclose reading a vocabulary data, comparing the additional word with the script data, or removing the additional word from the script data.

In the same field of speech recognition, GOULD teaches:

reading a vocabulary data having one or more words ("dictionary", column 15, line 15);

comparing the additional word with the script data ("if the text on the screen starting with the current word matches the indicated words, set the selection to text on the screen just compared against", column 13, lines 35-38, where "text on the screen" is the additional word, and the "indicated words" is the script data);

removing the additional word from the script data if the additional word is included in the script data ("if words are selected on the screen, delete the words which are selected", column 13, lines 48-49, where the word "selected on the screen" is the additional word that was compared with the script data).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to delete the words in the new sentence provided by ESQUERRA, in view of NEWMAN, in the manner of GOULD in order to filter out words which may be problematic for training (see PITRELLI et al., section 2.1, listed on form PTO-892).

36. Regarding **claim 11**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), each word having one or more of the set of one or more phonemes ("N is the number of phones in a sentence", section 3.1, paragraph 1).

However, ESQUERRA, in view of NEWMAN, does not disclose reading vocabulary data, comparing words to vocabulary data, or returning an error message.

In the same field of speech recognition, GOULD teaches:

reading vocabulary data having one or more words ("dictionary", column 15, line 15);

comparing each word in the script data with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", column 15, lines 14-15); and

returning an error message if a word in the script data is not included in the vocabulary data ("if the word is not in the dictionary... display an 'unknown word' error to the user", column 15, lines 16-20).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the corpus of sentences provided by ESQUERRA, in view of NEWMAN, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

Regarding **claim 12**, ESQUERRA teaches counting each phoneme in each word in the script data ("units were counted to know whether they reach the minimum number of required repetitions", section 3.1, paragraph 1, where "units" refer to phonemes).

However, ESQUERRA, in view of NEWMAN, does not disclose comparing words to vocabulary data, returning an error message, or counting the phonemes if the word is in the vocabulary data.

In the same field of speech recognition, ESQUERRA in view of GOULD teach:

comparing each word in the script data with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", GOULD, column 15, lines 14-15);

returning an error message if a word in the script data is not included in the vocabulary data ("if the word is not in the dictionary... display an 'unknown word' error to the user", GOULD, column 15, lines 16-20); and

counting each phoneme in each word in the script data ("units were counted to know whether they reach the minimum number of required repetitions", ESQUERRA, section 3.1, paragraph 1) if a word in the script data is included in the vocabulary data ("remember these words as target words", column 15, line 21, where a word marked as a target word has further operations performed on it).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the corpus of sentences provided by ESQUERRA, in view of NEWMAN, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

37. Regarding **claim 16**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), and further comprising the steps of:

reading an additional word having one or more phonemes ("new sentences had to be written containing those allophones", section 3.1, paragraph 2); and

adding an additional word to the script data ("new sentences were added to the corpus", section 3.1, paragraph 4).

However ESQUERRA, in view of NEWMAN, does not disclose reading a vocabulary data, comparing the additional word to the vocabulary data, or adding the additional word if the additional word is included in the vocabulary data.

In the same field of speech recognition, ESQUERRA in view of GOULD teach: reading a vocabulary data having one or more words ("dictionary", GOULD, column 15, line 15);

comparing the additional word with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", GOULD, column 15, lines 14-15);

adding the additional word to the script data ("new sentences were added to the corpus", ESQUERRA, section 3.1, paragraph 4) if the additional word is included in the vocabulary data ("remember these words as target words", column 15, line 21, where a word marked as a target word has further operations performed on it).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the new sentence provided by ESQUERRA, in view of NEWMAN, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20).

38. Regarding **claim 17**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), and further comprising the step of:

reading an additional word having one or more phonemes ("new sentences had to be written containing those allophones", section 3.1, paragraph 2);

However ESQUERRA, in view of NEWMAN, does not disclose reading a vocabulary data, comparing the additional word with the script data, or removing the additional word from the script data.

In the same field of speech recognition, GOULD teaches:

reading a vocabulary data having one or more words ("dictionary", column 15, line 15);

comparing the additional word with the script data ("if the text on the screen starting with the current word matches the indicated words, set the selection to text on the screen just compared against", column 13, lines 35-38, where "text on the screen" is the additional word, and the "indicated words" is the script data);

removing the additional word from the script data if the additional word is included in the script data ("if words are selected on the screen, delete the words which are selected", column 13, lines 48-49, where the word "selected on the screen" is the additional word that was compared with the script data).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to delete the words in the new sentence provided by ESQUERRA, in view of NEWMAN, in the manner of GOULD in order to filter out words which may be problematic for training (see PITRELLI et al., section 2.1, listed on form PTO-892).

39. Regarding **claim 20**, ESQUERRA teaches that the script ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1), and wherein the tool is further programmed to read an additional word having one or more phonemes ("new sentences had to be written containing those allophones", section 3.1, paragraph 2), and add the additional word to the script data ("new sentences were added to the corpus", section 3.1, paragraph 4).

However ESQUERRA, in view of NEWMAN, does not disclose a tool that is programmed to read a vocabulary data having one or more words, and is also programmed to compare the additional word with the vocabulary data and add the additional word to the script data if the additional word is included in the vocabulary data, and is also programmed to compare the additional word with the script and remove the additional word from the script data if the additional word is included in the script data.

In the same field of speech recognition, ESQUERRA in view of GOULD teach a tool that is programmed to read a vocabulary data having one or more words ("dictionary", GOULD, column 15, line 15), and is also programmed to compare the additional word with the vocabulary data ("for each word in the buffer, look the word up in the dictionary", GOULD, column 15, lines 14-15) and add the additional word to the script data ("new sentences were added to the corpus", ESQUERRA, section 3.1, paragraph 4) if the additional word is included in the vocabulary data ("remember these words as target words", column 15, line 21, where a word marked as a target word has

further operations performed on it), and is also programmed to compare the additional word with the script ("if the text on the screen starting with the current word matches the indicated words, set the selection to text on the screen just compared against", GOULD, column 13, lines 35-38, where "text on the screen" is the additional word, and the "indicated words" is the script data) and remove the additional word from the script data if the additional word is included in the script data ("if words are selected on the screen, delete the words which are selected", GOULD, column 13, lines 48-49, where the word "selected on the screen" is the additional word that was compared with the script data).

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to check the words in the new sentence provided by ESQUERRA, in view of NEWMAN, in the manner of GOULD in order to ensure that a speech model can be obtained for each word (see GOULD, column 15, lines 14-20) and to filter out words which may be problematic for training (see PITRELLI et al., section 2.1, listed on form PTO-892).

40. **Claims 9 and 18** are rejected under 35 U.S.C. 103(a) as being unpatentable over ESQUERRA et al. ("Design of a Phonetic Corpus for Speech Recognition in Catalan"), in view of NEWMAN, and further in view of Department of Psychology, University of Essex ("Phoneme Search"), hereinafter referred to as ESSEX.

41. Regarding **claim 9**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1).

However, ESQUERRA, in view of NEWMAN, does not disclose reading a vocabulary data, reading a set of one or more desired phonemes, searching the vocabulary data for one or more words having the set of one or more desired phonemes, or generating a report of one or more additional words having the set of one or more desired phonemes.

In the same field of phonetic evaluation, ESSEX teaches:

reading a vocabulary data having one or more words ("word database", see header);

reading a set of one or more desired phonemes (three different phonemes may be selected with the pull-down menus);

searching the vocabulary data for one or more words having the set of one or more desired phonemes ("search for words which contain the following phonemes");

generating a report of one or more additional words having the set of one or more desired phonemes ("generates a list of words"), if the one or more additional words having the set of one or more desired phonemes are included in the vocabulary data (see "Phoneme Search Results").

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made use the phoneme search engine of ESSEX with the

corpus design of ESQUERRA, in view of NEWMAN, in order to find words containing "missing units" (ESSEX, section 5, paragraph 1).

42. Regarding **claim 18**, ESQUERRA teaches that the script data ("corpus of sentences", section 3.1, paragraph 1) includes one or more words ("sentences between 10 and 40 letters were selected", section 3.1, paragraph 1).

However, ESQUERRA, in view of NEWMAN, does not disclose reading a vocabulary data, reading a set of one or more desired phonemes, searching the vocabulary data for one or more words having the set of one or more desired phonemes, or generating a report of one or more additional words having the set of one or more desired phonemes.

In the same field of phonetic evaluation, ESSEX teaches:

reading a vocabulary data having one or more words ("word database", see header);

reading a set of one or more desired phonemes (three different phonemes may be selected with the pull-down menus);

searching the vocabulary data for one or more words having the set of one or more desired phonemes ("search for words which contain the following phonemes");

generating a report of one or more additional words having the set of one or more desired phonemes ("generates a list of words"), if the one or more additional words having the set of one or more desired phonemes are included in the vocabulary data (see "Phoneme Search Results").

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made use the phoneme search engine of ESSEX with the corpus design of ESQUERRA, in view of NEWMAN, in order to find words containing "missing units" (ESSEX, section 5, paragraph 1).

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to ERIC YEN whose telephone number is (571)272-4249. The examiner can normally be reached on M-F 7:30-4:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached on 571-272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Application/Control Number: 10/712,445
Art Unit: 2626

Page 47

EY 10/31/08

/Patrick N. Edouard/
Supervisory Patent Examiner, Art Unit 2626